Physica A 459 (2016) 24–31

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Combining fuzzy logic and eigenvector centrality measure in social network analysis

Fereshteh-Azadi Parand *, Hossein Rahimi, Mohsen Gorzin

Math and Computer Science Department, Allameh Tabataba'i University, Tehran, Iran

HIGHLIGHTS

- A combined algorithm of fuzzy inference system and eigenvector centrality is proposed.
- Social interactions are measured by different factors with different weights.
- The influencing factors in a social network are used to weight the friendship strength using the fuzzy logic.
- The most influential person is calculated using eigenvector centrality after feeding it with the fuzzy logic results.
- The method is applied on large data sets such as Facebook, Epinions, and Slashdot-zoo website.

ARTICLE INFO

Article history:
Received 14 June 2015
Received in revised form 25 February 2016
Available online 28 April 2016

Keywords:
Social network analysis
Eigenvector centrality
Fuzzy inference system

ABSTRACT

The rapid growth of social networks use has made a great platform to present different services, increasing beneficiary of services and business profit. Therefore considering different levels of member activities in these networks, finding highly active members who can have the influence on the choice and the role of other members of the community is one the most important and challenging issues in recent years. These nodes that usually have a high number of relations with a lot of quality interactions are called influential nodes. There are various types of methods and measures presented to find these nodes. Among all the measures, centrality is the one that identifies various types of influential nodes in a network. Here we define four different factors which affect the strength of a relationship. A fuzzy inference system calculates the strength of each relation, creates a crisp matrix in which the corresponding elements identify the strength of each relation, and using this matrix eigenvector measure calculates the most influential node. Applying our suggested method resulted in choosing a more realistic central node with consideration of the strength of all friendships.

1. Introduction

Social network members interact and communicate in a self-organizing globally coherent pattern that appears in and makes up the system [1,2]. These patterns become more apparent as network size increases. A global network analysis of highly related networks is not feasible due to variety and quantity of information that causes them to be uninformative. The structure of relationships between social entities can also be examined by social network analysis techniques. Large things such as persons, groups, organizations, websites or scholarly publications can be encountered as entities of a social network. Considering the growth of social networks applications of new types and methods of marketing are introduced. In Ref. [3]

* Corresponding author.
E-mail address: parand@atu.ac.ir (F.-A. Parand).

http://dx.doi.org/10.1016/j.physa.2016.03.079

the importance of content marketing in social networks such as Facebook is described. It mentions that popular contents shared among the members can guide to intelligent advertisement and marketing for specific goods which can result in better sales and business profit.

Different fields of study like sociology utilize the obtained information from graphs by applying centrality measures. These sets of information include the relative importance of nodes and edges. Eigenvector centrality is an example of such method that uses the eigenvectors of the adjacency matrix of the network and determines most frequently visited nodes. Some of the formally established measures of centrality are betweenness, eigenvector, degree, closeness, and Katz centrality [4]. Generally, the purpose or objective of the analysis determines the type of chosen centrality measure [5,2]. There are different studies in the literature applying these methods analyzing social networks.

In some networks node explicitly states the group memberships. In Ref. [6] 230 of such real-world networks were studied. Based on the study, a methodology is presented to compare and evaluate the difference of structural definitions of network communities that correspond to ground-truth communities. Afterward, their sensitivity, robustness, and performance in identifying the ground-truth are examined. In the subject of the semantic web and the level of trust of each source, properties which merge such trusts, and a class of functions are discussed and defined in Ref. [7]. In this work, the functions are applied to data from Epinions and BibServ [8].

The stability of betweenness centrality (BC) is evaluated in Ref. [9]. A metric is used to measure the importance of the vertices in the network and introduces a group testing algorithm. The results of this study show how the ranking of the vertices changes as the networks are perturbed and the algorithm can correctly identify the high valued BC vertices of stable networks. A parametric fuzzy closeness measure which allows relaxation of the condition of all other nodes is presented in Ref. [10]. This measure is defined for unweighted networks and evaluations on real and exemplary networks indicate that new information is provided by the fuzzy measure for closeness centrality in networks that are not provided by the classical measures. This measure is more robust to observation errors in the network. An approximation of betweenness centrality is studied and defined in Ref. [11]. The purpose is to build a predictive model of social networks. The methodology describes a bounded distance approximation of betweenness centrality designed for implementation within a parallel architecture. In Ref. [12], a network flow topology based on multi-dimensional variation is presented. The dimensions are trajectories that traffic may follow and the method of spread. Measures of centrality are matched to the kinds of flows that they are appropriate for. One of the most recent works on real social networks is illustrated in Ref. [13]. The authors consider signed variants of global network characteristics such as the clustering coefficient, node-level characteristics, and link-level characteristics on a technology news website called Slashdot. The relations between members are described as friends and foes based on the positive or negative endorsement. Eigenvectors of adjacency matrices are misapplied to asymmetric networks in which some positions are unchosen. An alternative measure of centrality is suggested in Ref. [14] for these networks that equal an eigenvector. A new formulation for node-centrality is presented in Ref. [15]. The results of this study show that some property satisfaction of eigenvector centrality measure depends on normalization. The bootstrap sampling procedures are used in Ref. [16] in order to determine how sampling affects the stability of different network centrality measures. Some of the well-known measures such as high-degree, betweenness, closeness, eigenvector, PageRank are presented and evaluated with ten data sets in Ref. [17]. According to the results, new centrality measure called DegreeDistance is presented. This measure chooses high-degree seeds in an appropriate distance from each other. Results indicate that some measures are more stable than others and that stability is also a function of network and study properties. Real eigenvector and eigenvalue of a real matrix are discussed in Ref. [18]. In this work, Tian extended the real eigenvector of a real matrix to fuzzy eigenvectors. In this work, we combine the fuzzy logic and eigenvector degree centrality measures and present a model which results in many realistic inferences. The rest of the paper organized as follows; Section 2 describes and compares different centrality measures. We illustrate the new model in Section 3. Section 4 talks the experimental results and finally Section 5 draws a conclusion.

2. Methods

Network analysis takes advantage of different types of methods. These methods assess different aspects of the network and try to clarify some properties about entities in the network. Centrality methods are indicators of the most important vertices within a graph. Some of the interesting applications include identifying the most influential person or favorite content in a social network, a spread of disease, and new marketing opportunities. Centrality concepts were first developed in social network analysis, and the defined terms are reflecting the sociological origin of such networks in most a proper way [14]. Here we introduce different types of centralizers. The degree centrality is defined as the number of links incident upon a node. The degree can be interpreted in terms of the opportunity of the node to catch the flowing interactions in the network. The degree centrality of a vertex v for graph G is defined as below, where V and E indicate vertices and edges, respectively.

$$C_D(v) = \deg(v).$$

Closeness centrality is another method used in connected graphs. The natural distance between all pairs of nodes is metric used in this method. This metric is based on the length of the shortest paths. The closeness of a node is defined by Bavelas